Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 2980–2984, May 1987
Medical Sciences

cDNA sequence of human transforming gene *hst* and identification of the coding sequence required for transforming activity

(oncogene/transformation/human stomach cancer)

MASANORI TAIRA, TERUHIKO YOSHIDA, KIYOSHI MIYAGAWA, HIROMI SAKAMOTO*, MASAAKI TERADA, AND TAKASHI SUGIMURA

Genetics Division, National Cancer Center Research Institute, Tsukiji 5-chome, Chuo-ku, Tokyo 104, Japan

Contributed by Takashi Sugimura, December 22, 1986

ABSTRACT **The *hst* gene was originally identified as a transforming gene in DNAs from human stomach cancers and from a noncancerous portion of stomach mucosa by DNA-mediated transfection assay using NIH3T3 cells. cDNA clones of *hst* were isolated from the cDNA library constructed from poly(A)$^+$ RNA of a secondary transformant induced by the DNA from a stomach cancer. The sequence analysis of the *hst* cDNA revealed the presence of two open reading frames. When this cDNA was inserted into an expression vector containing the simian virus 40 promoter, it efficiently induced the transformation of NIH3T3 cells upon transfection. It was found that one of the reading frames, which coded for 206 amino acids, was responsible for the transforming activity.**

The DNA transfection assay using NIH3T3 mouse cells has been successfully used to detect transforming genes in various cancers. The majority of the transforming genes detected in this assay have been shown to be one of three closely related *ras* genes activated by a point mutation (for review, see refs. 1 and 2).

We have shown previously that 3 out of 58 samples of DNA taken from human stomach mucosa possessed transforming activity; one was taken from a primary stomach cancer, the second from a noncancerous portion of stomach mucosa from the same patient, and the third was from a lymph node metastasis of stomach cancer from a different patient. Portions of the gene that were expressed in the transformants have been cloned. The acquisition of transforming activity by all three samples of DNA was shown to be caused by the same transforming gene, *hst* (for *h*uman *st*omach cancer) (3).

In the present paper, we report the isolation of an *hst* cDNA clone that was shown to have an efficient transforming activity in a focus-forming assay when it was inserted into an expression vector. The characterization of this clone allows us to predict that a 206-amino acid protein product is responsible for the transforming activity.

MATERIALS AND METHODS

Culture and Cells. NIH3T3 cells and the transformants were cultured as described previously (3). T361 and T363 cells were the primary transformants induced by DNA samples from a stomach cancer (no. 361) and a noncancerous mucosa of the stomach (no. 363) from the same patient, respectively. The primary transformant T51 cells were induced by a DNA sample from a lymph node metastasis of a stomach cancer taken from the other patient (no. 51) (3). The transfection of the cloned cDNA into the NIH3T3 cells was carried out as described previously (3) after mixing it with 30 μg of DNA from the NIH3T3 cells.

Library Construction. Total RNA was isolated from T361-2nd-1 cells, a secondary transformant induced by the transfection of the human stomach cancer DNA, no. 361, using the CsCl/guanidinium isothiocyanate method (ref. 3; ref. 4, p. 196). Poly(A)$^+$ RNA was obtained by passing the total RNA through an oligo(dT)-cellulose column (ref. 4, pp. 197–198), and a cDNA library was constructed from the poly(A)$^+$ RNA by using the bacteriophage vector λgt10 (5, 6). The recombinant phages were plated onto *Escherichia coli* strain C600hfl$^+$. Phage DNA was transferred to duplicate nitrocellulose filters and denatured. Duplicate filters were hybridized with a nick-translated ^{32}P-labeled genomic *Sst* I–*Eco*RI 0.78-kilobase-pair (kbp) DNA fragment of the *hst* gene (3) at 42°C for 16–24 hr in a hybridization solution containing 50% (vol/vol) formamide, 0.65 M NaCl, 0.1 M Pipes–NaOH at pH 6.8, 5 mM EDTA, 0.1% NaDodSO$_4$, 10% dextran sulfate, 5× Denhardt's solution, 100 μg of denatured salmon testis DNA per ml, and ^{32}P-labeled probe at 2–5 × 10^6 cpm/ml. The filters were washed in a buffer containing 2× NaCl/Cit (1× NaCl/Cit is 0.15 M NaCl/0.015 M sodium citrate), 20 mM sodium phosphate buffer at pH 7.0, 0.06% sodium pyrophosphate, and 0.05% NaDodSO$_4$ at 52°C (ref. 4, pp. 326–328).

Nucleic Acid Hybridization. Southern blot hybridization analysis of DNA and blot hybridization analysis of RNA were performed as described previously (ref. 3; ref. 4, pp. 387–389).

DNA Sequence Analysis. *hst* cDNA was inserted into the *Eco*RI site of M13mp18 phage and the overlapping subclones were generated by the stepwise deletion method (7). Both strands of cDNA were sequenced in full by the dideoxy chain termination method (8). Deoxy-7-deazaguanosine triphosphate was used in place of dGTP to sequence a G+C-rich portion of the gene (9). The Maxam–Gilbert method (10) was used to confirm the results obtained with the dideoxy chain termination method. Homology search among known DNA and protein sequences was performed in the GenBank† nucleic acid database and the National Biomedical Research Foundation‡ protein data base, respectively, with a VAX/VMS computer using the algorithms developed by Wilbur and Lipman (11).

RESULTS

cDNA Cloning. The *Sst* I–*Eco*RI 0.78-kbp DNA fragment, a portion of the *hst* genomic sequence previously shown to be expressed in the transformants, was used to screen the cDNA library constructed from poly(A)$^+$ RNA in T361-2nd-1 cells.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: SV40, simian virus 40.

*To whom reprint requests should be addressed.

†National Institutes of Health (1986) Genetic Sequence Databank: GenBank (Research Systems Div., Bolt, Beranek, and Newman, Cambridge, MA), Tape Release 42.0.

‡Protein Identification Resource (1986) Protein Sequence Database (Natl. Biomed. Res. Found., Washington, DC), Release 9.0.

By using this probe, a major transcript of 3.3-kilobase (kb) was detected in the poly(A)$^+$ RNA from T361-2nd-1 cells (3). The screening of 110,000 plaques with this probe resulted in the isolation of 4 clones, and the restriction enzyme mapping demonstrated that the four clones were overlapping (data not shown). The size of the two 3.2-kbp cDNA clones was almost identical to that of the major transcript of the *hst* gene in the T361-2nd-1 cells, and the restriction maps for the two cDNA clones were identical. One of the 3.2-kbp cDNA clones was designated λCT361-b3. Blot hybridization analyses of poly(A)$^+$ RNA from T361-2nd-1 cells showed that this cDNA insert of λCT361-b3 hybridized strongly to the 3.3-kb RNA but not to the RNA from the NIH3T3 cells (Fig. 1). This result showed that the 3.2-kbp cDNA represented a near-full-length cDNA for the major transcript of the *hst* gene. When the cDNA clone was used as a probe, Southern blot hybridization analysis of DNA samples from the primary transformants, T361, T363, and T51 cells, showed strongly hybridized bands of human origin in addition to the weakly hybridized mouse-type fragment (data not shown). These results confirmed our previous findings, showing that all the transformants induced by three different samples of DNAs contained the same transforming gene, although there were some rearrangements at the time of transfection (3). The transcriptional orientation was determined by blot hybridization analysis of RNA using a single-strand probe of cDNA that was cloned in M13mp18 phage. It was established that the left-hand side of the restriction map shown in Fig. 2*A* was the 5′ end of the mRNA for *hst*.

cDNA Sequence and Open Reading Frames. The nucleotide sequence for the 3149 nucleotides of the cDNA insert of λCT361-b3 was determined (Fig. 2*B*). The G+C content of the 5′ region (from nucleotide 1 to 600) was very high, being more than 75%, while that of the 3′ region (from nucleotide 2000 to 3149) was less than 40%. Stop codons, TAA, TAG, and TGA, were present mainly in the 3′ A+T-rich region. There were two open reading frames that started with the initiation codon and could code for more than 153 amino acids, ORF 1 from nucleotide 239 to 856 and ORF 2 from nucleotide 1218 to 1679 in the 5′ parts of the cDNA. At the 3′ end of the cDNA sequence, the AATAAA polyadenylylation signal and poly(dA) stretch were not present despite the oligo(dT)-primed synthesis of cDNA. The cDNA inserts in the other cDNA clones were sequenced at the 3′ ends but no poly(dA) stretch was found.

Transforming Sequences in cDNA. To prove that the cDNA clone contained a transforming sequence, the cDNA from the λCT361-b3 clone was inserted into an expression vector that had been constructed by removing the *neo* gene from pKOneo (12) (Fig. 3). The 3149-nucleotide sequence of the *hst* cDNA was inserted between the SV40 early promoter and the poly(A) addition site, and the resulting plasmid was designated pKOc1. pKOc1 had an efficient transforming activity of approximately 4000 foci per μg of DNA upon transfection of NIH3T3 cells. The transformants were tumorigenic to *nude* mice. In contrast, pKOc2, in which the orientation of the *hst* cDNA was reversed, had no transforming activity. The transforming efficiency of pKOc1 was approximately the same as that of activated c-Ha-*ras* cloned from T24 bladder carcinoma (13). To determine which open reading frame was responsible for the transforming activity, an insertion mutation and a deletion mutation were introduced into ORF 2 and ORF 1, respectively. pKOc11, in which ORF 2 was destroyed by the insertion of GATC, at the *Bgl* II site between positions 1415 and 1416 in ORF 2, had the same transforming efficiency as pKOc1. In contrast, pKOc12, in which 103 bp between *Sst* II sites was deleted from positions 159 to 261, including 23 bp from positions 239 to 261 of ORF 1, had no transforming activity. Furthermore, pKOc5, which contained the nucleotide sequence from positions 1 to 916 and included only ORF 1, had the same transforming efficiency as pKOc1 (Fig. 3). We concluded that ORF 1 was responsible for the transforming activity of the *hst* cDNA. The amino acid sequence in ORF 1 of the *hst* gene was deduced from the nucleotide sequence (Fig. 2). ORF 1 encodes a 206-amino acid polypeptide with a predicted molecular weight of 22,000. A homology search of the 3149-bp cDNA sequence was conducted in the GenBank nucleotide sequence library. Computer analyses of the entire 206-amino acid sequence of ORF 1 were performed with the National Biomedical Research Foundation protein data base and also with the amino acid sequences deduced from the GenBank nucleotide database. In no case was any significant homology found to the genes compiled in the data bases. These include *abl*, B-*lym*, *erbA*, *erbB-1*, *erbB-2*, *ets*, *fes*, *fgr*, *fms*, *fos*, *int-1*, *mos*, *myb*, *myc*, p53, *raf*, H-*ras*, K-*ras*, N-*ras*, *rel*, *sis*, *src*, epidermal growth factor, platelet-derived growth factor B chain, insulin-like growth factor I, insulin-like growth factor II, β-nerve growth factor, type-α transforming growth factor, type-β transforming growth factor, and gastrin-releasing peptide. Other recently reported oncogenes and growth factors, including *met* (14), N-*myc* (15), *trk* (16), *mas* (17), *ros* (18), *kit* (19), platelet-derived growth factor A chain (20), α-nerve growth factor (21), and B-cell growth factor II (22), had no strong homology with the ORF 1 of *hst* cDNA or with its deduced amino acid sequence. The DNA samples from all the transformants induced by nos. 361, 363, and 51 DNAs did not hybridize with *lca* probe (23).

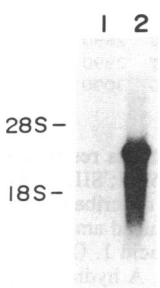


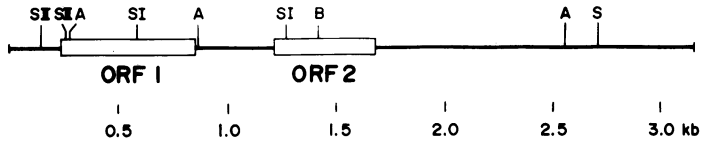
FIG. 1. Expression of *hst* gene in the transformant. Samples (10 μg) of poly(A)$^+$ RNA from NIH3T3 cells (lane 1) and from T361-2nd-1 cells (lane 2) were analyzed by blot hybridization using λcT361-b3 cDNA as a probe.

DISCUSSION

We have previously described the identification and cloning of a portion of the transforming gene *hst* (3). This *hst* gene was responsible for the acquisition of transforming activity in 3 out of 58 samples of DNA, which included a primary stomach cancer and a noncancerous portion of the stomach mucosa taken from the same patient and a lymph node metastasis of a stomach cancer from a different patient. In this paper, we have described the isolation of *hst* cDNA clones from a secondary transformant induced by DNA from the primary stomach cancer. Using *hst* cDNA as a probe, we were able to confirm our previous results that the *hst* gene was responsible for the acquisition of transforming activity in these three samples of DNA. The nucleotide sequence of *hst* cDNA has been determined. There are two open reading frames, ORF 1 and ORF 2, and ORF1 is responsible for the acquisition of the transforming activity. The biological significance of ORF 2 has yet to be determined.

The *hst* product responsible for the transforming activity was deduced from the nucleotide sequence of ORF 1 and was found to consist of 206 amino acids. The sequence surrounding the initiation codon, ATG, conforms well to the nucleotide sequence proposed for the start of eukaryotic translation (24). The codon usage in ORF 1 was consistent with that of animals (25). The analysis of the amino acid sequence predicts a signal peptide (26, 27), an N-linked glycosylation site (Asn-Xaa-Thr/Ser) (28), and four possible proteolytic

A



B

```
GCACTGCTCCTCAGAGTCCCAGCTCCAGCCGCGCGCTTTCCGCCCGGCTCGCCGCTCCATGCAGCCGGGGTAGAGCCCGGCGCCCGGGGGCCCCGTCGCT   100
TGCCTCCCGCACCTCCTCGGTTGCGCACTCCTGCCCGAGGTCGGCCGTGCGCTCCCGCGGGACGCCACAGGCGCAGCTCTGCCCCCCAGCTTCCCGGGCG   200

CACTGACCGCCTGACCGACGCACGGCCCTCGGGCCGGG  ATG TCG GGG CCC GGG ACG GCC GCG GTA GCG CTG CTC CCG GCG GTC      283
                                       Met Ser Gly Pro Gly Thr Ala Ala Val Ala Leu Leu Pro Ala Val
                                       1

CTG CTG GCC TTG CTG GCG CCC TGG GCG GGC CGA GGG GGC GCC GCC GCA CCC ACT GCA CCC AAC GGC ACG CTG GAG      358
Leu Leu Ala Leu Leu Ala Pro Trp Ala Gly Arg Gly Gly Ala Ala Ala Pro Thr Ala Pro Asn Gly Thr Leu Glu
                20                                                                  ★  ★  ★      40

GCC GAG CTG GAG CGC CGC TGG GAG AGC CTG GTG GCG CTC TCG TTG GCG CGC CTG CCG GTG GCA GCG CAG CCC AAG      433
Ala Glu Leu Glu Arg Arg Trp Glu Ser Leu Val Ala Leu Ser Leu Ala Arg Leu Pro Val Ala Ala Gln Pro Lys
                                                                                    60

GAG GCG GCC GTC CAG AGC GGC GCC GGC GAC TAC CTG CTG GGC ATC AAG CGG CTG CGG CGG CTC TAC TGC AAC GTG      508
Glu Ala Ala Val Gln Ser Gly Ala Gly Asp Tyr Leu Leu Gly Ile Lys Arg Leu Arg Arg Leu Tyr Cys Asn Val
                                                        80

GGC ATC GGC TTC CAC CTC CAG GCG CTC CCC GAC GGC CGC ATC GGC GGC GCG CAC GCG GAC ACC CGC GAC AGC CTG      583
Gly Ile Gly Phe His Leu Gln Ala Leu Pro Asp Gly Arg Ile Gly Gly Ala His Ala Asp Thr Arg Asp Ser Leu
                                100

CTG GAG CTC TCG CCC GTG GAG CGG GGC GTG GTG AGC ATC TTC GGC GTG GCC AGC CGG TTC TTC GTG GCC ATG AGC      658
Leu Glu Leu Ser Pro Val Glu Arg Gly Val Val Ser Ile Phe Gly Val Ala Ser Arg Phe Phe Val Ala Met Ser
                120                                                                 140

AGC AAG GGC AAG CTC TAT GGC TCG CCC TTC TTC ACC GAT GAG TGC ACG TTC AAG GAG ATT CTC CTT CCC AAC AAC      733
Ser Lys Gly Lys Leu Tyr Gly Ser Pro Phe Phe Thr Asp Glu Cys Thr Phe Lys Glu Ile Leu Leu Pro Asn Asn
                                                                    160

TAC AAC GCC TAC GAG TCC TAC AAG TAC CCC GGC ATG TTC ATC GCC CTG AGC AAG AAT GGG AAG ACC AAG AAG GGG      808
Tyr Asn Ala Tyr Glu Ser Tyr Lys Tyr Pro Gly Met Phe Ile Ala Leu Ser Lys Asn Gly Lys Thr Lys Lys Gly
                                        180

AAC CGA GTG TCG CCC ACC ATG AAG GTC ACC CAC TTC CTC CCC AGG CTG TGA  CCCTCCAGAGGACCCTTGCCTCAGCCTCGGG   890
Asn Arg Val Ser Pro Thr Met Lys Val Thr His Phe Leu Pro Arg Leu End
                                200

AAGCCCCTGGGAGGGCAGTGCCGAGAGTCACCTTGGTGCACTTTCTTCGGATGAAGAGTTTAATGCAAGAGTAGGTGTAAGATATTTAAATTAATTATTT   990
AAATGTGTATATATTGCCACCAAATTATTTATAGTTCTGCGGGTGTGTTTTTTAATTTTCTGGGGGGAAAAAAAGACAAAACAAAAAACCAACTCTGACT  1090
TTTCTGGTGCAACAGTGGAGAATCTTACCATTGGATTTCTTTAACTTGTCAAAAGTTGTCACGAGTGTGCTGCTATTCTGTGTTTTAAAAAAAGGTGACA  1190
TTGGATTCCGATGTCATCCCCTGTAGTATGGCGTGGAGCATCTCTGTCTGGAAAGGCCCGCCTGAGGCTTGGGCAGCCAGTTCAGGGAGCTCCCAGGCTT  1290
GGCTCTCGGCTAGCATCCTCAGAGGCCCACTCCCTTTGTGCCCTGTTGCTATTAATCGGGACATATCGGTTTACTTCGGGTACAGAAAGTGCGGTGTTGA  1390
AGTCCTCGCTGCCACTCTGTTTTTAGATCTGCCAAGACTGACCTTTGAACTTTCCTGTAGTCAATCTTCCTCGATCTACCAGATGGGAGAGACCCTTGGA  1490
CAACTTTATAAACTCCTGTTTGCCTTTTTTGGATCAGCGACAGCCCCCATCGCTGTGACTATTGGGGAAAAGACGAAGCTCTTTCATAAATTCCATGGAG  1590
AGGAATCAATATCCCACTGGAAGGCTAGAAATGGACAAGATAGTGTATTTGCAATCACAAACAAAACCCTAGTGATGAAAAATAATTTGTGATGGCAGAT  1690
GCTTCTGATGGTGTGATAGAATATGTTTTTGAAAACAAACCATCGAACCCCCCGCCCCACCCCCAAAACGGGCTTCCCTGTGTTTAGGGAGCTTTGGGCT  1790
AGAACTAGCTACGATTTTTAGGTGAAATGTCCTTGTAATTGTACAAAGCACTTGGTGCAGTGTTTGCGTGGAGCAGCCTGCTGCTTTCTGATGCATTCCC  1890
TGTTTAAGTGCGTTTAACATCTACCTCACAAGCCCTGAAACCCCAGGCAAAACCCACAGAAAGCTCATACCCGGTGCAGGAGTTTGCCATCCCAAGTGGC  1990
TTTTTTTCCATATGTAGCCAAAAAGGATTGCAGATAGCGTCGGTGCGTCCCATTCGAACCTTGTCACGTTTGAGCTATCTTTACCCTGTGATTTACTTTT  2090
AGTAAGGGTGATCATGGTGAAAATATTTGCAGACAGCTGTTACAGTACACTATATGGTCACCAAGTAACCTTATATTTTTCTTTATATATTTTACAAATG  2190
TAACCCCTGTCATTGAAGCAACCGTGGAAGAGGCAGGGTCGGTGATGTTTAAAAAAAGTTCCGAGGTGATGGCAAACATTTAATTTTAATGAATGACTTT  2290
TTAGAGTTTATACAAAATGACCTTAGCTTGCTACCAGAAATGCTCCGAATGTTTCGTCAAGACTTTAATACTCTCCTAGGATGTTTCTGAACTGTCTCCC  2390
GAATTAACTTTATGGGAGTCTACAGACAGCAAGACTGGAAAATCTGATTGGAGTTTTTGTCTTTCACATTCCTTTTGAAAACTCTTTGTTCGAATGCAAA  2490
TCATCGACTTAAAATACTATTCTTAACCAAGGCCTGGAAGAAAGAAGACACTTGCAAAGCCGCTAAGACAGGACCACACATCTTAAACTGCTGTTCCTAC  2590
CATGCACTAAACTGTTTTTAAGTTTTAAACCACACCCTAGGCTCCAGGAGTGTTCAGGAAAGATGGTGTTTGTAGGTCTCCATGCTGTTTGGCGTTGGGG  2690
GGTGTGGAGGGATCATCCGTCGACTTTCTGAATTTTAATGTATTCACTTAGTAACAAACCATGATTGTCTTAAATGCCTTAAATTATTATGAGATTTCTT  2790
GTCTCAGAGCCCAATCAGATTGTCAGGAATTAACATGTGTTAGGTTTGATCACCCTTGACCACTTCTTATAGATATTTCTTCAACAAATCATGTGTGATG  2890
CCTGTAGGAACACAACTGTACCTTTAAAATATTGTTTTCATATTGCTGTGATGGGGATTCGAGGTTCCTGTATGTGCCACTGTTTTCAGAATCTGTAGTT  2990
TTATACAGGTGCCGACCCTCGTTGTGATGTATGTGCTGTGCACATTGACATGCTGACCGACAATGATAAGCGTTTATCGTGTATAAAAAGACACCACTGG  3090
ACTGGATGTACACAACTGGGAAAGGAATTAAAAGCTATTAAAATTGTGCCTTGAAATGC  3149
```

FIG. 2. Nucleotide sequence and deduced amino acid sequence of *hst* cDNA. (*A*) *hst* cDNA is shown schematically with its restriction endonuclease maps and the two open reading frames, ORF 1 and ORF 2. Restriction sites: A, *Ava* II; B, *Bgl* II; S, *Sal* I; SI, *Sst* I; SII, *Sst* II. (*B*) Nucleotide sequence and deduced amino acid sequence of *hst* cDNA cloned in λCT361-b3. Sequences were obtained as described in the text and are shown from the left to the right in the 5′-to-3′ orientation. Nucleotides are numbered at the end of lines. The deduced amino acid sequence of ORF 1 is shown under the nucleotide sequence and is numbered by designating the first methionine as amino acid 1. ORF 2 is underlined with a solid line. In ORF 1, possible glycosylation (★) and dibasic proteolytic cleavage (〰) sites are indicated. A hydrophobic sequence with a property of core sequence for a possible signal peptide is underlined with a broken line.

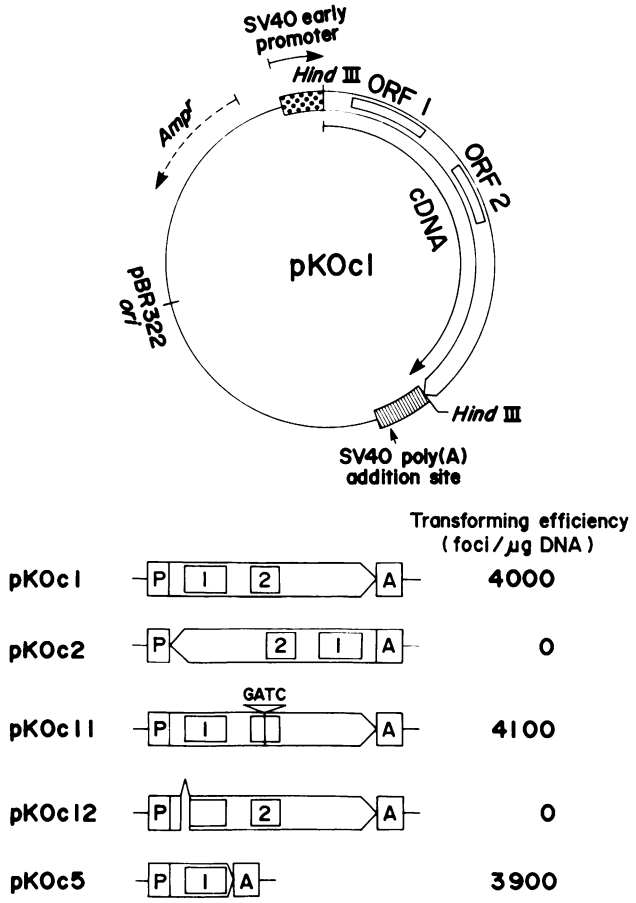


FIG. 3. Transforming activity of *hst* cDNA in an expression vector. (*Upper*) Detailed map of the vector. The cDNA insert that had been cloned at the *Eco*RI site of λgt10 was excised from λCT361-b3. This *Eco*RI fragment, to which *Hin*dIII linker was attached after treatment with *E. coli* DNA polymerase I Klenow fragment, was inserted into the *Hin*dIII site of pKOneo between the simian virus 40 (SV40) promoter and the polyadenylylation site, replacing the *neo* gene with the cDNA insert. pKOc1 is presented schematically at the top of *Lower* [P indicates promoter and A, poly(A) addition site]. It contained 3149 base pairs (bp) of cDNA in the 5′-to-3′ direction. pKOc2 contained cDNA in the reverse direction. In pKOc11, ORF 2 was destroyed by inserting GATC between nucleotides 1415 and 1416, while in pKOc12 nucleotides from positions 159 to 261 were deleted, resulting in the removal of 23 bp from nucleotide positions 239 to 261 in ORF 1. pKOc5 contained ORF 1 but not ORF 2. To construct pKOc5, a cDNA insert spanning from nucleotide positions 1 to 916 was obtained from an M13 subclone of λCT361-b3 and inserted at the *Hin*dIII site of pKOneo by using *Hin*dIII linker. These clones were transfected into NIH3T3 cells with carrier DNA and the foci were counted. The transforming activity of each gene construct is presented at the right in *Lower*.

sites consisting of dibasic amino acids (26), Arg-45 and Arg-46, Lys-81 and Arg-82, Arg-84 and Arg-85, and Lys-188 and Lys-189. It should be noted that Arg-Xaa-Arg/Lys-Arg are proteolytic sites for various growth factors such as platelet-derived growth factor A chain (20) and B chain (29), type-β transforming growth factor (30), and β-nerve growth factor (31). The sequence Arg-Leu-Arg-Arg was present at amino acid residues 82–85. The N-terminal region of the amino acid residues could be a signal peptide, containing a hydrophobic core sequence of amino acids from residues 9 to 20, which are common features of a signal peptide (27). There is no ATP (32) or GTP binding site (33), and the amino acid sequence is not homologous with either serine/threonine-specific or tyrosine-specific protein kinases (34). These results indicate that the *hst* gene belongs neither to the kinase gene family nor to the *ras* gene family. The presence of a possible signal peptide and of possible proteolytic sites suggest that the *hst* gene product with the molecular weight of approximately 22,000 may be a growth factor (20, 26, 27, 29–31, 35–37).

Although the *hst* gene product may be a growth factor with transforming activity, the exact nature of the product remains unknown. Knowledge of the amino acid structure deduced from the *hst* cDNA used as a molecular probe will certainly contribute to the elucidation of the role of the *hst* gene in the development of stomach cancer, which has the highest incidence of all cancers in the world (38).

Note Added in Proof. After submission of the manuscript, the National Biomedical Research Foundation protein data base was updated to Release 10.0. Homology search against this latest version of the data base revealed 42.3% homology between the amino acid residues 70–204 of ORF 1 of *hst* and the residues 6–141 of bovine basic fibroblast growth factor.

We thank Drs. M. Wigler and D. Hanahan for providing us with the recombinant DNA clones, Dr. K. Matsubara for providing us the *lca* probe, and Dr. A. Ito for helping us with computer analyses. This study was supported in part by a grant-in-aid from the Ministry of Health and Welfare for a Comprehensive 10-Year Strategy for Cancer Control, Japan. M. Taira, T.Y., and K.M. received Research Resident Fellowships from the Foundation for Promotion of Cancer Research, Japan.

1. Duesberg, P. H. (1985) *Science* **228,** 669–677.
2. Bishop, J. M. (1985) *Trends Genet.* **1,** 245–249.
3. Sakamoto, H., Mori, M., Taira, M., Yoshida, T., Matsukawa, S., Shimizu, K., Sekiguchi, M., Terada, M. & Sugimura, T. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 3997–4001.
4. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
5. Huynh, T. V., Young, R. A. & Davis, R. W. (1984) in *DNA Cloning: A Practical Approach*, ed. Glover, D. M. (IRL, Oxford), Vol. 1, pp. 49–78.
6. Gubler, U. & Hoffman, B. J. (1983) *Gene* **25,** 263–269.
7. Yanisch-Perron, C., Vieira, J. & Messing, J. (1985) *Gene* **33,** 103–119.
8. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
9. Mizusawa, S., Nishimura, S. & Seela, F. (1986) *Nucleic Acids Res.* **14,** 1319–1324.
10. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65,** 499–560.
11. Wilbur, W. J. & Lipman, D. J. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 726–730.
12. Fasano, O., Birnbaum, D., Edlund, L., Fogh, J. & Wigler, M. (1984) *Mol. Cell. Biol.* **4,** 1695–1705.
13. Fasano, O., Taparowsky, E., Fiddes, J., Wigler, M. & Goldfarb, M. (1983) *J. Mol. Appl. Genet.* **2,** 173–180.
14. Dean, M., Park, M., Le Beau, M. M., Robins, T. S., Diaz, M. O., Rowley, J. D., Blair, D. G. & Vande Woude, G. F. (1985) *Nature* **318,** 385–388.
15. Stanton, L. W., Schwab, M. & Bishop, J. M. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 1772–1776.
16. Martin-Zanca, D., Hughes, S. H. & Barbacid, M. (1986) *Nature (London)* **319,** 743–748.
17. Young, D., Waitches, G., Birchmeier, C., Fasano, O. & Wigler, M. (1986) *Cell* **45,** 711–719.
18. Matsushime, H., Wang, L.-H. & Shibuya, M. (1986) *Mol. Cell Biol.* **6,** 3000–3004.
19. Besmer, P., Murphy, J. E., George, P. C., Qiu, F., Bergold, P. J., Lederman, L., Snyder, H. W., Jr., Brodeur, D., Zuckerman, E. E. & Hardy, W. D. (1986) *Nature (London)* **320,** 415–421.
20. Betsholtz, C., Johnsson, A., Heldin, C.-H., Westermark, B., Lind, P., Urdea, M. S., Eddy, R., Shows, T. B., Philpott, K., Mellor, A. L., Knott, T. J. & Scott, J. (1986) *Nature (London)* **320,** 695–699.
21. Evans, B. A. & Richards, R. I. (1985) *EMBO J.* **4,** 133–138.
22. Kinashi, T., Harada, N., Severinson, E., Tanabe, T., Sideras, P., Konishi, M., Azuma, C., Tominaga, A., Bergstedt-Lind-

qvist, S., Takahashi, M., Matsuda, F., Yaoita, Y., Takatsu, K. & Honjo, T. (1986) *Nature (London)* **324,** 70–73.
23. Ochiya, T., Fujiyama, A., Fukushige, S., Hatada, I. & Matsubara, K. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 4993–4997.
24. Kozak, M. (1984) *Nucleic Acids Res.* **12,** 857–872.
25. Grantham, R., Gautier, C., Gouy, M., Jacobzone, M. & Mercier, R. (1981) *Nucleic Acids Res.* **9,** r43–r74.
26. Steiner, D. F., Quinn, P. S., Chan, S. J., Marsh, J. & Tager, H. S. (1980) *Ann. N.Y. Acad. Sci.* **343,** 1–16.
27. Perlman, D. & Nalvorson, H. D. (1983) *J. Mol. Biol.* **167,** 391–409.
28. Wagh, P. V. & Bahl, O. P. (1981) *CRC Crit. Rev. Biochem.* **10,** 307–377.
29. Rao, C. D., Igarashi, H., Chiu, I.-M., Robbins, K. C. & Aaronson, S. A. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 2392–2396.
30. Derynck, R., Jarrett, J. A., Chen, E. Y., Eaton, D. H., Bell, J. R., Assoian, R. K., Roberts, A. B., Sporn, M. B. & Goeddel, D. V. (1985) *Nature (London)* **316,** 701–705.
31. Meier, R., Becker-Andre, M., Gotz, R., Heumann, R., Shaw, A. & Thoenen, H. (1986) *EMBO J.* **5,** 1489–1493.
32. Hannink, M. & Donoghue, D. J. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 7894–7898.
33. Tanabe, T., Nukada, T., Nishikawa, Y., Sugimoto, K., Suzuki, H., Takahashi, H., Noda, M., Haga, T., Ichiyama, A., Kangawa, K., Minamino, N., Matsuo, H. & Numa, S. (1985) *Nature (London)* **315,** 242–245.
34. Parker, P. J., Coussens, L., Totty, N., Rhee, L., Young, S., Chen, E., Stabel, S., Waterfield, M. D. & Ullrich, A. (1986) *Science* **233,** 853–859.
35. Lee, D. C., Rose, T. M., Webb, N. R. & Todaro, G. J. (1985) *Nature (London)* **313,** 489–491.
36. Bell, G. I., Merryweather, J. P., Sanchez-Pescador, R., Stempien, M. M., Priestley, L., Scott, J. & Rall, L. B. (1984) *Nature (London)* **310,** 775–777.
37. Abraham, J. A., Mergia, A., Whang, J. L., Tumolo, A., Friedman, J., Hjerrild, K. A., Gospodarowicz, D. & Fiddes, J. C. (1986) *Science* **233,** 545–548.
38. Parkin, D. M., Stjernsward, J. & Muir, C. S. (1984) *Bull. W. H. O.* **62,** 163–182